Results and questions on a nonlinear approximation approach for solving high-dimensional partial differential equations

C. Le Bris[1], T. Lelièvre [1] & Y. Maday [2]

[1] CERMICS, École des Ponts,
6 & 8, avenue Blaise Pascal, 77455 Marne-La-Vallée Cedex 2, FRANCE, and
INRIA Rocquencourt, MICMAC project-team,
Domaine de Voluceau, B.P. 105, 78153 Le Chesnay Cedex, FRANCE
{lebris,lelievre}@cermics.enpc.fr

[2] Laboratoire J.L.-Lions, Université Pierre et Marie Curie,
Boite courrier 187, F-75252 Paris, FRANCE
maday@ann.jussieu.fr

November 4, 2008

Abstract

We investigate mathematically a nonlinear approximation type approach recently introduced in [1] to solve high dimensional partial differential equations. We show the link between the approach and the *greedy algorithms* of approximation theory studied *e.g.* in [4]. On the prototypical case of the Poisson equation, we show that a variational version of the approach, based on minimization of energies, converges. On the other hand, we show various theoretical and numerical difficulties arising with the non variational version of the approach, consisting of simply solving the first order optimality equations of the problem. Several unsolved issues are indicated in order to motivate further research.

1 Introduction

Our purpose here is to investigate mathematically a numerical approach recently introduced in [1] to solve high dimensional partial differential equations.

The approach is a nonlinear approximation type approach that consists in expanding the solution of the equation in tensor products of functions sequentially determined as the iterations of the algorithm proceed. The original motivation of the approach is the wish of its authors to solve high-dimensional Fokker-Planck type equations arising in the modelling of complex fluids. Reportedly, the approach performs well in this case, and, in addition, extends to a large variety of partial differential equations, static or time-dependent, linear or nonlinear, elliptic or parabolic, involving self-adjoint or non self- adjoint operators provided the data enjoy some appropriate separation property with respect to the different coordinates (this property is made precise in Remark 1 below). We refer the reader to [1] for more details.

In the present contribution focused on mathematical analysis, we restrict ourselves to the simplest possible case, namely the solution of the Poisson equation set with Dirichlet homogeneous boundary conditions on a two dimensional parallelepipedic domain $\Omega =$

$\Omega_x \times \Omega_y$ with $\Omega_x \subset \mathbb{R}$ and $\Omega_y \subset \mathbb{R}$ bounded. In short, the approach under consideration then determines the solution u to

$$-\Delta u(x,y) = f(x,y) \tag{1}$$

as a sum

$$u(x,y) = \sum_{n\geq 1} r_n(x)\, s_n(y), \tag{2}$$

by iteratively determining functions $r_n(x)$, $s_n(y)$, $n \geq 1$ such that for all n, $r_n(x)\, s_n(y)$ is the best approximation (in a sense to be made precise below) of the solution $v(x,y)$ to $-\Delta v(x,y) = f(x,y) + \Delta\left(\sum_{k\leq n-1} r_k(x)\, s_k(y)\right)$ in terms of one single tensor product $r(x)s(y)$. We show that it is possible to give a sound mathematical ground to the approach *provided* we consider a variational form of the approach that manipulates minimizers of energies instead of solutions to equations. In order to reformulate the approach in such a variational setting, our arguments thus crucially exploit the fact that the Laplace operator is self-adjoint. It is to be already emphasized that, because of the nonlinearity of the tensor product expansion (2), the variational form of the approach is *not* equivalent to the form (1)-(2) (which is exactly the Euler-Lagrange equations associated to the energy considered in the variational approach). Our analysis therefore does not apply to the actual implementation of the method as described in [1]. At present time, we do not know how to extend our arguments to cover the practical situation, even in the simple case of the Poisson problem. The consideration of some particular pathological cases, theoretically and numerically, shows that the appropriate mathematical setting is unclear. Likewise, it is unclear to us how to provide a mathematical foundation of the approach for non variational situations, such as an equation involving a differential operator that is not self-adjoint.

On the other hand, the analysis provided here straightforwardly extends to the case of a N-dimensional Poisson problem with $N \geq 3$ (unless explicitly mentioned). Likewise, our analysis extends to the case of elliptic linear partial differential equations set on a cylinder in $\mathbb{R}^N$, with appropriate boundary conditions. The only, although substantial, difficulty that may appear when the dimension N grows is the algorithmic complexity of the approach, since a set of N coupled non-linear equations has to be solved (see Remark 2). At least, the number of unknowns involved in the systems to be solved does not grow exponentially, as it would be the case for a naive approach (like for a finite differences method on a tensorized grid). This is not the purpose of the present article to further elaborate on this.

Our article is organized as follows. Section 2 introduces the approach. The variational version of the approach (along with a relaxed variant of it) is described in Section 2.1. Elementary properties follow in Sections 2.2 and 2.3. The non variational version is presented in Section 2.4. In Section 3 we show the convergence of the variational approach and give an estimate of the rate of convergence. Our arguments immediately follow from standard arguments of the literature of *nonlinear approximation theory*, and especially from those of [4]. The particular approach under consideration is indeed closely related to the so-called *greedy algorithms* introduced in approximation theory. We refer to [2, 3, 8] for some relevant contributions, among many. The purpose of Section 4 is to return to the original non variational formulation of the approach. For illustration, we first consider the case when the Laplace operator $-\Delta$ in (1) is replaced by the identity operator. The approach then reduces to the determination of the *Singular Value Decomposition* (also called *rank-one decomposition*) of the right-hand side f. This simple situation allows one to understand various difficulties inherent to the non variational formulation of the approach. We then discuss the actual case of the Laplace operator, and present some intriguing numerical

experiments, in particular when a non-symmetric term (namely there an advection term) is added.

As will be clear from the sequel, our current mathematical understanding of the numerical approach is rather incomplete. Our results do not cover real practice. Some ingredients from the literature of nonlinear approximation theory nevertheless already allow for understanding some basics of the approach. It is the hope of the authors that, laying some groundwork, the present contribution will sparkle some interest among the experts, and allows in a not too far future for a complete understanding of the mathematical nature of the approach. Should the need arise, it will also indicate possible improvements of the approach so that it is rigorously founded mathematically and, eventually, performs even better that the currently existing reports seemingly show.

Acknowledgments: The authors wish to thank A. Ammar and F. Chinesta for introducing them to their series of works initiated in [1], A. Cohen for stimulating discussions, and A. Lozinski for pointing out reference [4]. This work was completed while the first author (CLB) was a long-term visitor at the Institute for Mathematics and its Applications, Minneapolis. The hospitality of this institution is gratefully acknowledged.

2 Presentation of the algorithms

Consider a function $f \in L^2(\Omega)$ where $\Omega = \Omega_x \times \Omega_y$ with $\Omega_x \subset \mathbb{R}$ and $\Omega_y \subset \mathbb{R}$ two bounded domains. To fix ideas, one may take $\Omega_x = \Omega_y = (0,1)$. Consider on Ω the following homogeneous Dirichlet problem:

$$\text{Find } g \in H^1_0(\Omega) \text{ such that } \left\{ \begin{array}{rl} -\Delta g = f & \text{in } \Omega, \\ g = 0 & \text{on } \partial\Omega. \end{array} \right. \tag{3}$$

It is well known that solving (3) is equivalent to solving the variational problem:

$$\text{Find } g \in H^1_0(\Omega) \text{ such that } g = \arg \min_{u \in H^1_0(\Omega)} \left(\frac{1}{2} \int_\Omega |\nabla u|^2 - \int_\Omega f u \right). \tag{4}$$

In the following, for any function $u \in H^1_0(\Omega)$, we denote

$$\mathcal{E}(u) = \frac{1}{2} \int_\Omega |\nabla u|^2 - \int_\Omega f u. \tag{5}$$

Notice that

$$\mathcal{E}(u) = \frac{1}{2} \int_\Omega |\nabla (u-g)|^2 - \frac{1}{2} \int_\Omega |\nabla g|^2 \tag{6}$$

where g is defined by (3), so that minimizing $\mathcal{E}$ is equivalent to minimizing $\int_\Omega |\nabla(u-g)|^2$ with respect to u. We endow the functional space $H^1_0(\Omega)$ with the scalar product:

$$\langle u, v \rangle = \int_\Omega \nabla u \cdot \nabla v,$$

and the associated norm

$$\|u\|^2 = \langle u, u \rangle = \int_\Omega |\nabla u|^2.$$

2.1 Two algorithms

We now introduce two algorithms to solve (3). The first algorithm is the *Pure Greedy Algorithm*: set $f_0 = f$, and at iteration $n \geq 1$,

1. Find $r_n \in H_0^1(\Omega_x)$ and $s_n \in H_0^1(\Omega_y)$ such that

$$(r_n, s_n) = \arg \min_{(r,s)\in H_0^1(\Omega_x)\times H_0^1(\Omega_y)} \left(\frac{1}{2}\int_\Omega |\nabla(r\otimes s)|^2 - \int_\Omega f_{n-1}\, r\otimes s \right). \tag{7}$$

2. Set $f_n = f_{n-1} + \Delta(r_n \otimes s_n)$.

3. If $\|f_n\|_{H^{-1}(\Omega)} \geq \varepsilon$, proceed to iteration $n+1$. Otherwise, stop.

Throughout this article, we denote by $r\otimes s$ the tensor product: $r\otimes s(x,y) = r(x)s(y)$. Notice that

$$f_n = f + \Delta\left(\sum_{k=1}^n r_k \otimes s_k\right).$$

The fonction f_n belongs to $H^{-1}(\Omega)$ and the tensor product $r\otimes s$ is in $H_0^1(\Omega)$ if $r \in H_0^1(\Omega_x)$ and $s \in H_0^1(\Omega_y)$ (see Lemma 1 below), so that the integral $\int_\Omega f_{n-1}\, r\otimes s$ in (7) is well defined.

A variant of this algorithm is the *Orthogonal Greedy Algorithm*: set $f_0^o = f$, and at iteration $n \geq 1$,

1. Find $r_n^o \in H_0^1(\Omega_x)$ and $s_n^o \in H_0^1(\Omega_y)$ such that

$$(r_n^o, s_n^o) = \arg \min_{(r,s)\in H_0^1(\Omega_x)\times H_0^1(\Omega_y)} \left(\frac{1}{2}\int_\Omega |\nabla(r\otimes s)|^2 - \int_\Omega f_{n-1}^o\, r\otimes s \right). \tag{8}$$

2. Solve the following Galerkin problem on the basis $(r_1^o\otimes s_1^o, \ldots, r_n^o\otimes s_n^o)$: find $(\alpha_1,\ldots,\alpha_n) \in \mathbb{R}^n$ such that

$$(\alpha_1,\ldots,\alpha_n) = \arg \min_{(\beta_1,\ldots,\beta_n)\in\mathbb{R}^n} \left(\frac{1}{2}\int_\Omega \left|\nabla\left(\sum_{k=1}^n \beta_k r_k^o \otimes s_k^o\right)\right|^2 - \int_\Omega f \sum_{k=1}^n \beta_k r_k^o\otimes s_k^o \right). \tag{9}$$

3. Set $f_n^o = f + \Delta\left(\sum_{k=1}^n \alpha_k r_k^o \otimes s_k^o\right)$.

4. If $\|f_n^o\|_{H^{-1}(\Omega)} \geq \varepsilon$, proceed to iteration $n+1$. Otherwise, stop.

Let us also introduce g_n satisfying the Dirichlet problem:

$$\begin{cases} -\Delta g_n = f_n & \text{in } \Omega, \\ \quad g_n = 0 & \text{on } \partial\Omega. \end{cases} \tag{10}$$

Notice that

$$g_n = g_{n-1} - r_n \otimes s_n. \tag{11}$$

so that $g_n = g - \sum_{k=1}^n r_k \otimes s_k$. Likewise, we introduce $g_n^o = g - \sum_{k=1}^n r_k^o \otimes s_k^o$, which satisfies $-\Delta g_n^o = f_n^o$ in Ω and $g_n^o = 0$ on $\partial\Omega$. Proving the convergence of the algorithms amounts to proving that g_n and g_n^o converge to 0.

The terminology *Pure Greedy Algorithm* and *Orthogonal Greedy Algorithm* is borrowed from approximation theory (see [2, 3, 4, 8]). Such algorithms have been introduced in a

more general framework, namely for an arbitrary Hilbert space and an arbitrary set of functions (not only tensor products). Recall for consistency that, in short, the purpose of such nonlinear approximations techniques is to find the best possible approximation of a given function as a sum of elements of a prescribed *dictionary*. The latter does not need to have a vectorial structure. In the present case, the dictionary is the set of simple products $r(x)s(y)$ for r varying in $H^1_0(\Omega_x)$ and s varying in $H^1_0(\Omega_y)$ (All this will be formalized with the introduction of the space $\mathcal{L}^1$ in Section 3 below). The metric chosen to define the approximation is the natural metric induced by the differential operator, here the H^1 norm. The algorithm proposed by Ammar *et al.* [1] is actually related to the Orthogonal Greedy Algorithm: it consists in replacing the optimization procedure (8) by the associated Euler-Lagrange equations. We shall give details on this in Section 2.3 below. For the moment, we concentrate ourselves on the variational algorithms above.

2.2 The iterations are well defined

We will need the following three lemmas.

Lemma 1 *For any measurable functions $r : \Omega_x \to \mathbb{R}$ and $s : \Omega_y \to \mathbb{R}$ such that $r \otimes s \neq 0$*

$$r \otimes s \in H^1_0(\Omega) \iff r \in H^1_0(\Omega_x) \text{ and } s \in H^1_0(\Omega_y).$$

Lemma 2 *Let $T \in \mathcal{D}'(\Omega)$ be a distribution such that, for any functions $(\phi, \psi) \in \mathcal{C}^\infty_c(\Omega_x) \times \mathcal{C}^\infty_c(\Omega_y)$,*

$$(T, \phi \otimes \psi)_{(\mathcal{D}'(\Omega),\mathcal{D}(\Omega))} = 0$$

then $T = 0$ in $\mathcal{D}'(\Omega)$. Moreover, for any two sequences of distributions $R_n \in \mathcal{D}'(\Omega_x)$ and $S_n \in \mathcal{D}'(\Omega_y)$ such that $\lim_{n\to\infty} R_n = R$ in $\mathcal{D}'(\Omega_x)$ and $\lim_{n\to\infty} S_n = S$ in $\mathcal{D}'(\Omega_y)$, $\lim_{n\to\infty} R_n \otimes S_n = R \otimes S$ in $\mathcal{D}'(\Omega)$.

Lemma 3 *Let us consider a function $f \in L^2(\Omega)$. If $f \neq 0$, then $\exists (r, s) \in H^1_0(\Omega_x) \times H^1_0(\Omega_y)$ such that*

$$\mathcal{E}(r \otimes s) < 0,$$

where $\mathcal{E}$ is defined by (5).

Lemma 2 is well-known in distribution theory. We now provide for consistency a short proof of Lemmas 1 and 3, respectively.

Proof of Lemma 1 Notice that

$$\int_\Omega |\nabla(r \otimes s)|^2 = \int_{\Omega_x} |r'|^2 \int_{\Omega_y} |s|^2 + \int_{\Omega_x} |r|^2 \int_{\Omega_y} |s'|^2$$

where $'$ denotes henceforth the differentiation with respect to a one-dimensional argument. Thus, it is clear that if $r \in H^1_0(\Omega_x)$ and $s \in H^1_0(\Omega_y)$, then $r \otimes s \in H^1_0(\Omega)$. Now, when $r \otimes s \in H^1_0(\Omega)$, we have $\int_{\Omega_x} |r'|^2 \int_{\Omega_y} |s|^2 < \infty$ and $\int_{\Omega_x} |r|^2 \int_{\Omega_y} |s'|^2 < \infty$. This implies $r \in H^1_0(\Omega_x)$ and $s \in H^1_0(\Omega_y)$, since $r \neq 0$ and $s \neq 0$. $\diamondsuit$

Proof of Lemma 3 Fix $f \in L^2(\Omega)$ and assume that for all $(r, s) \in H^1_0(\Omega_x) \times H^1_0(\Omega_y)$, $\mathcal{E}(r \otimes s) \geq 0$. Then, for a fixed $(r, s) \in H^1_0(\Omega_x) \times H^1_0(\Omega_y)$, we have, for all $\epsilon \in \mathbb{R}$,

$$\frac{\epsilon^2}{2} \int |\nabla(r \otimes s)|^2 \geq \epsilon \int f r \otimes s.$$

By letting $\epsilon \to 0$, this shows that $f \in \{r \otimes s, (r, s) \in L^2(\Omega_x) \times L^2(\Omega_y)\}^\perp$ which implies $f = 0$ (by Lemma 2) and concludes the proof. $\diamondsuit$

The above lemmas allow us to prove.

Proposition 1 *For each n, there exists a solution to problems* (7) *and* (8).

Proof. Without loss of generality, we may only argue on problem (7) and assume that $n = 1$ and $f_0 = f \neq 0$. First, using (6), it is clear that

$$\frac{1}{2}\int_\Omega |\nabla(r\otimes s)|^2 - \int_\Omega f\, r\otimes s = \frac{1}{2}\int_\Omega |\nabla(r\otimes s - g)|^2 - \frac{1}{2}\int_\Omega |\nabla g|^2$$
$$\geq -\frac{1}{2}\int_\Omega |\nabla g|^2.$$

Thus, we can introduce $m = \inf_{(r,s)\in H^1_0(\Omega_x)\times H^1_0(\Omega_y)} \left(\frac{1}{2}\int_\Omega |\nabla(r\otimes s)|^2 - \int_\Omega f\, r\otimes s\right)$ and a minimizing sequence (r^k, s^k) such that $\lim_{k\to\infty} \mathcal{E}(r^k\otimes s^k) = m$. Notice that we may suppose, again without loss of generality (up to a multiplication of s^k by a constant), that

$$\int_\Omega |r^k|^2 = 1.$$

Since $\mathcal{E}(u) \geq \frac{1}{4}\int_\Omega |\nabla u|^2 - \int_\Omega |\nabla g|^2$, the sequence $(r^k\otimes s^k)$ is bounded in $H^1_0(\Omega)$: there exists some $C > 0$ such that, for all $k \geq 1$,

$$\int_{\Omega_x} |(r^k)'|^2 \int_{\Omega_y} |s^k|^2 + \int_{\Omega_x} |r^k|^2 \int_{\Omega_y} |(s^k)'|^2 \leq C. \tag{12}$$

From this we deduce the existence of $w \in H^1_0(\Omega)$, $r \in L^2(\Omega_x)$ and $s \in H^1_0(\Omega_y)$ such that (up to the extraction of a subsequence):

- $r^k\otimes s^k$ converges to w weakly in $H^1_0(\Omega)$, and strongly in $L^2(\Omega)$,
- r^k converges to r weakly in $L^2(\Omega_x)$,
- s^k converges to s weakly in $H^1_0(\Omega_y)$, and strongly in $L^2(\Omega_y)$.

Since $r^k\otimes s^k$ converges to w weakly in $H^1_0(\Omega)$ and $\mathcal{E}$ is convex and continuous, we have $\mathcal{E}(w) \leq \liminf_{k\to\infty} \mathcal{E}(r^k\otimes s^k)$. This yields $\mathcal{E}(w) \leq m$. Moreover, by Lemma 3, we know $m < 0$. Therefore,

$$\mathcal{E}(w) < 0. \tag{13}$$

The convergences $r^k \to r$ and $s^k \to s$ in the distributional sense imply the convergence $r^k\otimes s^k \to r\otimes s$ in the distributional sense (see Lemma 2), and therefore $w = r\otimes s$. Thus, if $w \neq 0$, Lemma 1 concludes the proof, showing that indeed $r \in H^1_0(\Omega_x)$. Now, we cannot have $w = 0$, since this would imply $\mathcal{E}(w) = 0$, which would contradict (13). This concludes the proof. ♢

The optimization step (9) admits also a solution by standard arguments and we therefore have proven:

Lemma 4 *At each iteration of the Pure Greedy Algorithm, problem (7) admits (at least) a minimizer (r_n, s_n). Likewise, at each iteration of the Orthogonal Greedy Algorithm, problem (8) admits (at least) a minimizer (r^o_n, s^o_n).*

It is important to note that, in either case, uniqueness of the iterate is unclear. Throughout the text, we will thus be refering to *the* functions (r_n, s_n) (resp. (r^o_n, s^o_n)) although we do not know whether they are unique. However, our arguments and results are valid for *any such* functions.

2.3 Euler-Lagrange equations

Our purpose is now to derive the Euler-Lagrange equations of the problems considered, along with other important properties of the sequences (r_n, s_n) and (r_n^o, s_n^o). We only state the results for (r_n, s_n). Similar properties hold for (r_n^o, s_n^o), replacing f_n and g_n by f_n^o and g_n^o.

The first order optimality conditions write:

Proposition 2 *The functions $(r_n, s_n) \in H_0^1(\Omega_x) \times H_0^1(\Omega_y)$ satisfying (7) are such that: for any functions $(r, s) \in H_0^1(\Omega_x) \times H_0^1(\Omega_y)$*

$$\int_\Omega \nabla(r_n \otimes s_n) \cdot \nabla(r_n \otimes s + r \otimes s_n) = \int_\Omega f_{n-1}(r_n \otimes s + r \otimes s_n). \tag{14}$$

This can be written equivalently as

$$\begin{cases} -\left(\displaystyle\int_{\Omega_y} |s_n|^2\right) r_n'' + \left(\displaystyle\int_{\Omega_y} |s_n'|^2\right) = r_n \displaystyle\int_{\Omega_y} f_{n-1}\, s_n, \\ \\ -\left(\displaystyle\int_{\Omega_x} |r_n|^2\right) s_n'' + \left(\displaystyle\int_{\Omega_x} |r_n'|^2\right) = s_n \displaystyle\int_{\Omega_x} f_{n-1}\, r_n, \end{cases} \tag{15}$$

or, in terms of g_n, as:

$$\langle g_n, (r \otimes s_n + r_n \otimes s)\rangle = 0. \tag{16}$$

Proof. Equation (14) is obtained differentiating (7). Namely, for any $(r, s) \in H_0^1(\Omega_x) \times H_0^1(\Omega_y)$ and any $\epsilon \in \mathbb{R}$, we have

$$\begin{aligned} &\frac{1}{2}\int_\Omega |\nabla\left((r_n + \epsilon r) \otimes (s_n + \epsilon s)\right)|^2 - \int_\Omega f_{n-1}\,(r_n + \epsilon r) \otimes (s_n + \epsilon s) \\ &\qquad \geq \frac{1}{2}\int_\Omega |\nabla(r_n \otimes s_n)|^2 - \int_\Omega f_{n-1}\, r_n \otimes s_n. \end{aligned} \tag{17}$$

It holds:

$$\begin{aligned} &\frac{1}{2}\int_\Omega |\nabla\left((r_n + \epsilon r) \otimes (s_n + \epsilon s)\right)|^2 - \int_\Omega f_{n-1}\,(r_n + \epsilon r) \otimes (s_n + \epsilon s) \\ &\quad = \frac{1}{2}\int_\Omega |\nabla(r_n \otimes s_n) + \epsilon\nabla(r \otimes s_n + r_n \otimes s) + \epsilon^2\nabla(r \otimes s)|^2 - \int_\Omega f_{n-1}\,(r_n + \epsilon r) \otimes (s_n + \epsilon s) \\ &\quad = \frac{1}{2}\int_\Omega |\nabla(r_n \otimes s_n)|^2 - \int_\Omega f_{n-1}\, r_n \otimes s_n \\ &\qquad + \epsilon\left(\int_\Omega \nabla(r_n \otimes s_n) \cdot \nabla(r \otimes s_n + r_n \otimes s) - \int_\Omega f_{n-1}(r_n \otimes s + r \otimes s_n)\right) \\ &\qquad + \epsilon^2\left(\frac{1}{2}\int_\Omega |\nabla(r \otimes s_n + r_n \otimes s)|^2 + \int_\Omega \nabla(r_n \otimes s_n) \cdot \nabla(r \otimes s) - \int_\Omega f_{n-1} r \otimes s\right) + O(\epsilon^3) \\ &\quad = \frac{1}{2}\int_\Omega |\nabla(r_n \otimes s_n)|^2 - \int_\Omega f_{n-1}\, r_n \otimes s_n + \epsilon I_1 + \epsilon^2 I_2 + O(\epsilon^3). \end{aligned}$$

Using (17), we get, for any $\epsilon \in \mathbb{R}$,

$$\epsilon I_1 + \epsilon^2 I_2 + O(\epsilon^3) \geq 0, \tag{18}$$

which implies that I_1 is zero, that is, (14).

Equation (15) is the strong formulation of (14). On the other hand, (16) is an immediate consequence of the following simple computations:

$$\begin{aligned}
\langle g_n, (r \otimes s_n + r_n \otimes s)\rangle &= \langle g_{n-1} - r_n \otimes s_n, (r \otimes s_n + r_n \otimes s)\rangle \\
&= \int_\Omega \nabla(g_{n-1} - r_n \otimes s_n) \cdot \nabla(r \otimes s_n + r_n \otimes s) \\
&= -\int_\Omega \Delta g_{n-1}(r \otimes s_n + r_n \otimes s) - \int_\Omega \nabla(r_n \otimes s_n) \cdot \nabla(r \otimes s_n + r_n \otimes s) \\
&= 0,
\end{aligned}$$

since $-\Delta g_{n-1} = f_{n-1}$ in Ω and $g_{n-1} = 0$ on $\partial\Omega$. $\diamondsuit$

Note that, taking $r = r_n$ and $s = 0$ in the Euler-Lagrange equations (16) yields

$$\langle r_n \otimes s_n, g_{n-1}\rangle = \|r_n \otimes s_n\|^2, \tag{19}$$

since $g_n = g_{n-1} - r_n \otimes s_n$. This will be useful below.

Let us now state two other properties of (r_n, s_n). The second order optimality conditions write:

Lemma 5 *The functions $(r_n, s_n) \in H^1_0(\Omega_x) \times H^1_0(\Omega_y)$ satisfying (7) are such that: for any functions $(r, s) \in H^1_0(\Omega_x) \times H^1_0(\Omega_y)$*

$$\frac{1}{2}\int_\Omega |\nabla(r \otimes s_n + r_n \otimes s)|^2 + \int_\Omega \nabla(r_n \otimes s_n) \cdot \nabla(r \otimes s) - \int_\Omega f_{n-1} r \otimes s \geq 0, \tag{20}$$

which is equivalent to: for any functions $(r, s) \in H^1_0(\Omega_x) \times H^1_0(\Omega_y)$

$$\left(\int_\Omega \nabla(r_n \otimes s_n - g_n) \cdot \nabla(r \otimes s)\right)^2 \leq \int_\Omega |\nabla(r \otimes s_n)|^2 \int_\Omega |\nabla(r_n \otimes s)|^2. \tag{21}$$

Proof. Returning to Equation (18), we see that $I_1 = 0$ and $I_2 \geq 0$, which is exactly (20). For any $\lambda \in \mathbb{R}$, taking $(\lambda r, s)$ as a test function in (20) shows

$$\frac{1}{2}\int_\Omega |\lambda\nabla(r \otimes s_n) + \nabla(r_n \otimes s)|^2 + \int_\Omega \lambda\nabla(r_n \otimes s_n) \cdot \nabla(r \otimes s) - \int_\Omega f_{n-1}\lambda r \otimes s \geq 0.$$

This equivalently reads

$$\begin{aligned}
&\frac{\lambda^2}{2}\int_\Omega |\nabla(r \otimes s_n)|^2 + \lambda\left(\int_\Omega (\nabla(r \otimes s_n) \cdot \nabla(r_n \otimes s) + \nabla(r_n \otimes s_n) \cdot \nabla(r \otimes s)) - \int_\Omega f_{n-1} r \otimes s\right) \\
&\quad + \frac{1}{2}\int_\Omega |\nabla(r_n \otimes s)|^2 \geq 0,
\end{aligned}$$

hence

$$\begin{aligned}
&\left(\int_\Omega (\nabla(r \otimes s_n) \cdot \nabla(r_n \otimes s) + \nabla(r_n \otimes s_n) \cdot \nabla(r \otimes s)) - \int_\Omega f_{n-1} r \otimes s\right)^2 \\
&\qquad \leq \int_\Omega |\nabla(r \otimes s_n)|^2 \int_\Omega |\nabla(r_n \otimes s)|^2.
\end{aligned}$$

This yields (21). $\diamondsuit$

We will also need the following optimality property of (r_n, s_n):

Lemma 6 *The functions* (r_n, s_n) *satisfying* (7) *are such that:* $\forall (r,s) \in H_0^1(\Omega_x) \times H_0^1(\Omega_y)$

$$\|r_n \otimes s_n\| = \frac{\langle r_n \otimes s_n, g_{n-1}\rangle}{\|r_n \otimes s_n\|} \geq \frac{\langle r \otimes s, g_{n-1}\rangle}{\|r \otimes s\|}.$$

Proof. We may assume without loss of generality that $n = 1$. The first equality is (19). To prove the inequality, let us introduce the supremum:

$$M = \sup_{(u,v)\in H_0^1(\Omega_x)\times H_0^1(\Omega_y), \|u\otimes v\|=1} \langle u \otimes v, g\rangle.$$

Using (19), we have

$$\|r_1 \otimes s_1\| = \frac{\langle r_1 \otimes s_1, g\rangle}{\|r_1 \otimes s_1\|} \leq M, \tag{22}$$

by definition of M. On the other hand, consider $(u^k, v^k)_{k\geq 0}$ a maximizing sequence associated to the supremum M: $\|u^k \otimes v^k\| = 1$ and $\lim_{k\to\infty}\langle u^k \otimes v^k, g\rangle = M$. We have, using (7), for all $k \geq 0$,

$$\begin{aligned}\|g - r_1 \otimes s_1\|^2 &\leq \|g - \langle g, u^k \otimes v^k\rangle\, u^k \otimes v^k\|^2 \\ &= \|g\|^2 - \langle g, u^k \otimes v^k\rangle^2,\end{aligned}$$

and, letting $k \to \infty$,

$$\|g - r_1 \otimes s_1\|^2 \leq \|g\|^2 - M^2. \tag{23}$$

Combining (22) and (23), we get

$$\begin{aligned}\|g - r_1 \otimes s_1\|^2 &\leq \|g\|^2 - M^2 \\ &\leq \|g\|^2 - \|r_1 \otimes s_1\|^2 \\ &= \|g\|^2 - 2\langle g, r_1 \otimes s_1\rangle + \|r_1 \otimes s_1\|^2 \\ &= \|g - r_1 \otimes s_1\|^2\end{aligned}$$

so that all the inequalities are actually equalities. By using the fact that, by (22), $M \geq 0$, we thus have

$$M = \|r_1 \otimes s_1\| = \frac{\langle r_1 \otimes s_1, g\rangle}{\|r_1 \otimes s_1\|}.$$

This concludes the proof. $\diamondsuit$

2.4 Some preliminary remarks on the non variational approach implemented

Before we get to the proof of the convergence of the approach in the next section, let us conclude Section 2 by some comments that relates the theoretical framework developed here to the practice.

It is important to already note, although we will return to this in Section 4 below, that the Euler-Lagrange equation is indeed the form of the algorithm manipulated in practice by the authors of [1]. The above variational setting is somewhat difficult to implement in practice. It requires to solve for the minimizers of (7) (and (8) respectively), which can be extremely demanding computationally. In their implementation of the approach (developed independently from the above nonlinear approximation theoretic framework), Ammar *et al.* therefore propose to search for the iterate (r_n, s_n) (and respectively (r_n^o, s_n^o)) not as a minimizer to optimization problems (7) and (8), but as a solution to the associated Euler-Lagrange equations (first order optimality conditions). The Pure Greedy algorithm is thus replaced by: set $f_0 = f$, and at iteration $n \geq 1$,

1. `Find` $r_n \in H_0^1(\Omega_x)$ `and` $s_n \in H_0^1(\Omega_y)$ `such that, for all functions` $(r,s) \in H_0^1(\Omega_x) \times H_0^1(\Omega_y)$`,` (14) `(or its equivalent form` (15)`) holds.`
2. `Set` $f_n = f_{n-1} + \Delta(r_n \otimes s_n)$`.`
3. `If` $\|f_n\|_{H^{-1}(\Omega)} \geq \varepsilon$`, proceed to iteration` $n+1$`. Otherwise, stop.`

The Orthogonal Greedy Algorithm is modified likewise.

As already explained in the introduction, and in sharp contrast to the situation encountered for linear problems, being a solution to the Euler-Lagrange equation does not guarantee being a minimizer in this nonlinear framework. We will point out difficulties originating from this in Section 4.

In addition to the above theoretical difficulty, and in fact somehow entangled to it, we have to mention that of course, the Euler-Lagrange equations (15), as a nonlinear system, need to be solved iteratively. In [1], a simple fixed point procedure is employed: choose $(r_n^0, s_n^0) \in H_0^1(\Omega_x) \times H_0^1(\Omega_y)$ and, at iteration $k \geq 0$, compute $(r_n^k, s_n^k) \in H_0^1(\Omega_x) \times H_0^1(\Omega_y)$ solution to:

$$\left\{ \begin{aligned} & -\int_{\Omega_y} |s_n^k|^2 (r_n^{k+1})'' + \int_{\Omega_y} |(s_n^k)'|^2 r_n^{k+1} = \int_{\Omega_y} f_{n-1} s_n^k, \\ & -\int_{\Omega_x} |r_n^{k+1}|^2 (s_n^{k+1})'' + \int_{\Omega_x} |(r_n^{k+1})'|^2 s_n^{k+1} = \int_{\Omega_x} f_{n-1} r_n^{k+1}, \end{aligned} \right. \tag{24}$$

until convergence is reached. We will also discuss below the convergence properties of this procedure on simple examples.

Remark 1 *In practice (bearing in mind that the approach has been designed to solve high-dimensional problems), in order for the right-hand side terms in* (24) *to be computable, the function f needs to be expressed as a sum of tensor products. Otherwise, computing high dimensional integrals would be necessary, and this is a task of the same computational complexity as the original Poisson problem. The function f thus needs to enjoy some appropriate separation property with respect to the different coordinates.*

If f is not given in such a form, it may be possible to first apply the Singular Value Decomposition algorithm to get a good estimate of f as a sum of tensor products (see Section 4.1).

Remark 2 *In dimension $N \geq 2$ (on a parallelepipedic domain $\Omega = \Omega_{x_1} \times \ldots \times \Omega_{x_N}$), the Euler-Lagrange equations* (14) *become: find functions $(r_n^1, \ldots, r_n^N) \in H_0^1(\Omega_{x_1}) \times \ldots \times H_0^1(\Omega_{x_N})$ such that: for any functions $(r^1, \ldots, r^N) \in H_0^1(\Omega_{x_1}) \times \ldots \times H_0^1(\Omega_{x_N})$,*

$$\begin{aligned} & \int_\Omega \nabla(r_n^1 \otimes \ldots \otimes r_n^N) \cdot \sum_{k=1}^N \nabla(r_n^1 \otimes \ldots \otimes r_n^{k-1} \otimes r^k \otimes r_n^{k+1} \otimes \ldots \otimes r_n^N) \\ & \quad = \int_\Omega f_{n-1} \sum_{k=1}^N (r_n^1 \otimes \ldots \otimes r_n^{k-1} \otimes r^k \otimes r_n^{k+1} \otimes \ldots \otimes r_n^N). \end{aligned} \tag{25}$$

This is a nonlinear system of N equations, which only involves one-dimensional integrals by Fubini theorem, provided that the data f is expressed as a sum of tensor products (see Remark 1).

Remark 3 *We presented the algorithms without space discretization, which is required for the practical implementation. In practice, finite element spaces V_x^h (resp. V_y^h) are used to*

discretized $H_0^1(\Omega_x)$ (resp. $H_0^1(\Omega_y)$), where $h > 0$ denotes a space discretization parameter. The space discretized version of (14) *thus writes: find $(r_n^h, s_n^h) \in V_x^h \times V_y^h$ such that, for any functions $(r^h, s^h) \in V_x^h \times V_y^h$*

$$\int_\Omega \nabla(r_n^h \otimes s_n^h) \cdot \nabla(r_n^h \otimes s^h + r^h \otimes s_n^h) = \int_\Omega f_{n-1}^h (r_n^h \otimes s^h + r^h \otimes s_n^h). \tag{26}$$

3 Convergence

To start with, we prove that the approach converges. Then we will turn to the rate of convergence.

3.1 Convergence of the method

Theorem 1 [Pure Greedy Algorithm]

Consider the Pure Greedy Algorithm, and assume first that (r_n, s_n) satisfies the Euler-Lagrange equations (14)*. Denote by*

$$E_n = \frac{1}{2} \int_\Omega |\nabla(r_n \otimes s_n)|^2 - \int_\Omega f_{n-1}\, r_n \otimes s_n \tag{27}$$

the energy at iteration n. We have

$$\sum_n \int_\Omega |\nabla(r_n \otimes s_n)|^2 = -2 \sum_n E_n < \infty. \tag{28}$$

Assume in addition that (r_n, s_n) is a minimizer of (7)*. Then,*

$$\lim_{n\to\infty} g_n = 0 \textit{ in } H_0^1(\Omega). \tag{29}$$

Immediate consequences of (28) and (29) are

$$\lim_{n\to\infty} E_n = \lim_{n\to\infty} \|r_n \otimes s_n\| = 0,$$

and

$$\lim_{n\to\infty} f_n = 0 \text{ in } H^{-1}(\Omega).$$

Proof. Let us first suppose that (r_n, s_n) only satisfies the Euler-Lagrange equations (14). We notice that, using (16)

$$\begin{aligned} \|g_{n-1}\|^2 &= \|g_n + r_n \otimes s_n\|^2 \\ &= \|g_n\|^2 + \|r_n \otimes s_n\|^2. \end{aligned} \tag{30}$$

Thus, $\|g_n\|^2$ is a nonnegative non increasing sequence. Hence it converges. This implies that $\sum_n |\nabla(r_n \otimes s_n)|^2 < \infty$.

Next, notice that

$$\begin{aligned} E_n &= \frac{1}{2} \int_\Omega |\nabla(r_n \otimes s_n)|^2 - \int_\Omega f_{n-1}\, r_n \otimes s_n \\ &= \frac{1}{2} \int_\Omega |\nabla(r_n \otimes s_n)|^2 - \int_\Omega \nabla g_{n-1} \cdot \nabla(r_n \otimes s_n) \\ &= -\frac{1}{2} \int_\Omega |\nabla(r_n \otimes s_n)|^2, \end{aligned}$$

since by (19), $\int_\Omega \nabla g_{n-1} \cdot \nabla(r_n \otimes s_n) = \int_\Omega |\nabla(r_n \otimes s_n)|^2$. This proves the first part of the theorem. At this stage, we have only used that (r_n, s_n) satisfies the Euler-Lagrange equations (15).

To conclude that $\lim_{n\to\infty} f_n = 0$, we now need to assume that (r_n, s_n) indeed satisfies the minimization problem (7). We know that $\|g_n\|^2$ is a bounded sequence, and therefore, we may assume (up to the extraction of a subsequence) that g_n converges weakly in $H_0^1(\Omega)$ to some $g_\infty \in H_0^1(\Omega)$. For any $n \geq 1$ and for any functions $(r, s) \in H_0^1(\Omega_x) \times H_0^1(\Omega_y)$,

$$\frac{1}{2}\int_\Omega |\nabla(r \otimes s)|^2 - \int_\Omega \nabla g_{n-1} \cdot \nabla(r \otimes s) \geq E_n.$$

By passing to the limit this inequality, we have

$$\frac{1}{2}\int_\Omega |\nabla(r \otimes s)|^2 - \int_\Omega \nabla g_\infty \cdot \nabla(r \otimes s) \geq 0.$$

This implies that for any functions $(r, s) \in H_0^1(\Omega_x) \times H_0^1(\Omega_y)$,

$$\int_\Omega \nabla g_\infty \cdot \nabla(r \otimes s) = 0.$$

Thus, by Lemma 2, $-\Delta g_\infty = 0$ in the distributional sense, which, since $g_\infty \in H_0^1(\Omega)$, implies $g_\infty = 0$. This shows that there is only one possible limit for the subsequence g_n and thus that the whole sequence itself weakly converges to 0.

The convergence of g_n to 0 is actually strong in $H_0^1(\Omega)$. The argument we use here is taken from [5]. Using Lemma 6, we have: for any $n \geq m \geq 0$

$$\begin{aligned}
\|g_n - g_m\|^2 &= \|g_n\|^2 + \|g_m\|^2 - 2\left\langle g_n, \left(g_n + \sum_{k=m+1}^{n} r_k \otimes s_k\right)\right\rangle \\
&= \|g_n\|^2 + \|g_m\|^2 - 2\|g_n\|^2 - 2\sum_{k=m+1}^{n} \langle g_n, r_k \otimes s_k\rangle \\
&\leq -\|g_n\|^2 + \|g_m\|^2 + 2\sum_{k=m+1}^{n} \|r_k \otimes s_k\|\|r_{n+1} \otimes s_{n+1}\|.
\end{aligned}$$

Define $\phi(1) = 1$, $\phi(2) = \arg\min_{n>\phi(1)}\{\|r_n \otimes s_n\| \leq \|r_{\phi(1)} \otimes s_{\phi(1)}\|\}$, and, by induction,

$$\phi(k+1) = \arg\min_{n>\phi(k)}\{\|r_n \otimes s_n\| \leq \|r_{\phi(k)} \otimes s_{\phi(k)}\|\}.$$

Notice that $\lim_{k\to\infty} \phi(k) = \infty$ since $\lim_{k\to\infty} \|r_k \otimes s_k\| = 0$. For example, if $(\|r_k \otimes s_k\|)_{k\geq 1}$ is a decreasing sequence, then $\phi(k) = k$. Now, we have: for any $l \geq k \geq 0$

$$\begin{aligned}
\|g_{\phi(l)-1} - g_{\phi(k)-1}\|^2 &\leq -\|g_{\phi(l)-1}\|^2 + \|g_{\phi(k)-1}\|^2 + 2\sum_{i=\phi(k)}^{\phi(l)-1} \|r_i \otimes s_i\|\|r_{\phi(l)} \otimes s_{\phi(l)}\| \\
&\leq -\|g_{\phi(l)-1}\|^2 + \|g_{\phi(k)-1}\|^2 + 2\sum_{i=\phi(k)}^{\phi(l)-1} \|r_i \otimes s_i\|^2.
\end{aligned}$$

Since $\sum_{k\geq 1} \|r_k \otimes s_k\|^2 < \infty$ and $(\|g_n\|)_{n\geq 1}$ is converging, the previous inequality shows that the subsequence $(g_{\phi(l)-1})_{l\geq 0}$ is a Cauchy sequence, and therefore strongly converges

to 0 (recall it is already known that g_n weakly converges to 0). Since $\|g_n\|$ is itself a converging sequence, this shows that

$$\lim_{n\to\infty} \|g_n\| = 0.$$

$\diamondsuit$

A similar result holds for the Orthogonal Greedy Algorithm.

Theorem 2 [Orthogonal Greedy Algorithm]
Consider the Orthogonal Greedy Algorithm, and assume first that (r_n^o, s_n^o) only satisfies the Euler-Lagrange equations (14) *associated with* (8) *(thus with $(r_n, s_n, f_{n-1}) = (r_n^o, s_n^o, f_{n-1}^o)$ in* (14)*). Denote by*

$$E_n^o = \frac{1}{2}\int_\Omega |\nabla(r_n^o \otimes s_n^o)|^2 - \int_\Omega f_{n-1}^o\, r_n^o \otimes s_n^o \tag{31}$$

the energy at iteration n. We have

$$\sum_n \int_\Omega |\nabla(r_n^o \otimes s_n^o)|^2 = -2\sum_n E_n^o < \infty. \tag{32}$$

Assume in addition that (r_n^o, s_n^o) is indeed a minimizer to the optimization problem (8)*. Then,*

$$\lim_{n\to\infty} g_n^o = 0 \text{ in } H_0^1(\Omega). \tag{33}$$

Immediate consequences of (32) and (33) are

$$\lim_{n\to\infty} E_n^o = \lim_{n\to\infty} \|r_n^o \otimes s_n^o\| = 0,$$

and

$$\lim_{n\to\infty} f_n^o = 0 \text{ in } H^{-1}(\Omega).$$

Proof. Let us first assume that (r_n^o, s_n^o) only satisfies the Euler-Lagrange equations (14) (with $(r_n, s_n, f_{n-1}) = (r_n^o, s_n^o, f_{n-1}^o)$ in (14)). Notice that by (9) and (19):

$$\begin{aligned}\|g_n^o\|^2 &= \left\| g - \sum_{k=1}^n \alpha_k r_k^o \otimes s_k^o \right\|^2 \\ &\leq \|g_{n-1}^o - r_n^o \otimes s_n^o\|^2 \\ &= \|g_{n-1}^o\|^2 - \|r_n^o \otimes s_n^o\|^2.\end{aligned}$$

Thus, $\|g_n^o\|^2$ is a nonnegative non increasing sequence. Hence it converges. This implies that $\sum_{k\geq 1} \|r_k^o \otimes s_k^o\|^2 < \infty$, and proves the first part of the theorem, using the same arguments as in the proof of Theorem 1.

Let us now assume in addition that (r_n^o, s_n^o) is a minimizer to (8). For fixed r and s, we derive from (8):

$$-\frac{1}{2}\int_\Omega |\nabla(r_n^o\otimes s_n^o)|^2 = \frac{1}{2}\int_\Omega |\nabla(r_n^o\otimes s_n^o)|^2 - \int_\Omega f_{n-1}^o\, r_n^o\otimes s_n^o \leq \frac{1}{2}\int_\Omega |\nabla(r\otimes s)|^2 - \int_\Omega f_{n-1}^o\, r\otimes s.$$

Letting n go to infinity, and using the same arguments as in the proof of Theorem 1, this implies that g_n^o weakly converges to 0 in $H_0^1(\Omega)$. The proof of the strong convergence of g_n^o to zero is then easy since, using the Euler Lagrange equations associated to (9):

$$\|g_n^o\|^2 = \langle g_n^o, g\rangle,$$

and the right-hand side converges to 0. $\diamondsuit$

3.2 Rate of convergence of the method

We now present an estimate of the rate of convergence for both the Pure and the Orthogonal Greedy Algorithms. These results are borrowed from [4]. We begin by only citing the result for Pure Greedy Algorithm. On the other hand, with a view to showing the typical mathematical ingredients at play, we outline the proof of convergence of the Orthogonal Greedy Algorithm, contained in the original article [4].

We first need to introduce a functional space adapted to the convergence analysis (see [2, 4]).

Definition 1 *We define the $\mathcal{L}^1$ space as*

$$\mathcal{L}^1 = \left\{ g = \sum_{k\geq 0} c_k u_k \otimes v_k,\ \textit{where } u_k \in H_0^1(\Omega_x),\ v_k \in H_0^1(\Omega_y),\ \|u_k \otimes v_k\| = 1 \textit{ and } \sum_{k\geq 0} |c_k| < \infty \right\},$$

and we define the $\mathcal{L}^1$-norm as

$$\|g\|_{\mathcal{L}^1} = \inf \left\{ \sum_{k\geq 0} |c_k|, g = \sum_{k\geq 0} c_k u_k \otimes v_k,\ \textit{where } \|u_k \otimes v_k\| = 1 \right\},$$

for $g \in \mathcal{L}^1$.

The following properties may readily be established:

- The space $\mathcal{L}^1$ is a Banach space.
- The space $\mathcal{L}^1$ is continuously embedded in $H_0^1(\Omega)$.

Notice that, in the definition of $\mathcal{L}^1$, the function $g = \sum_{k\geq 0} c_k u_k \otimes v_k$ is indeed well defined in $H_0^1(\Omega)$ as a normally convergent series. This also shows that $\mathcal{L}^1 \subset H_0^1(\Omega)$, and this imbedding is continuous by the triangle inequality $\|\sum_{k\geq 0} c_k u_k \otimes v_k\| \leq \sum_{k\geq 0} |c_k|$.

We do not know if there exists a simple characterization of functions in $\mathcal{L}^1$. Let us however give simple examples of such functions.

Lemma 7 *For any $m > 2$, $H^m(\Omega) \cap H_0^1(\Omega) \subset \mathcal{L}^1$.*

Proof. Without loss of generality, consider the case $\Omega_x = \Omega_y = (0,1)$. Using the fact that $\{\phi_k \otimes \phi_l, k, l \geq 1\}$, where $\phi_k(x) = \sqrt{2}\sin(k\pi x)$, is an orthonormal basis of $L^2(\Omega)$, we can write any function $g \in L^2(\Omega)$ as the series $g = \sum_{k,l\geq 1} g_{k,l} \phi_k \otimes \phi_l$, where $g_{k,l} = \int_\Omega g\, \phi_k \otimes \phi_l$. It is well known that

$$g \in H_0^1(\Omega) \iff \sum_{k,l\geq 1} |g_{k,l}|^2 (k^2 + l^2) < \infty$$

and, more generally, for any $m \geq 1$,

$$g \in H^m(\Omega) \cap H_0^1(\Omega) \iff \sum_{k,l\geq 1} |g_{k,l}|^2 (k^2 + l^2)^m < \infty.$$

On the other hand,

$$\begin{aligned} \|g\|_{\mathcal{L}^1} &= \left\| \sum_{k,l\geq 1} g_{k,l} \phi_k \otimes \phi_l \right\|_{\mathcal{L}^1} \\ &= \left\| \sum_{k,l\geq 1} g_{k,l} \|\phi_k \otimes \phi_l\| \frac{\phi_k \otimes \phi_l}{\|\phi_k \otimes \phi_l\|} \right\|_{\mathcal{L}^1} \\ &\leq \sum_{k,l\geq 1} |g_{k,l}| \pi \sqrt{k^2 + l^2}, \end{aligned}$$

since $\|\phi_k \otimes \phi_l\| = \pi\sqrt{k^2+l^2}$. Thus, by the Hölder inequality, we have, for any $m > 2$, if $g \in H^m(\Omega) \cap H_0^1(\Omega)$,

$$
\begin{aligned}
\|g\|_{\mathcal{L}^1} &\leq \pi \sum_{k,l \geq 1} |g_{k,l}|(k^2+l^2)^{m/2}(k^2+l^2)^{(1-m)/2} \\
&\leq \pi \left(\sum_{k,l\geq 1} |g_{k,l}|^2 (k^2+l^2)^m \right)^{1/2} \left(\sum_{k,l \geq 1} (k^2+l^2)^{1-m} \right)^{1/2} \\
&< \infty,
\end{aligned}
$$

since $\sum_{k,l\geq 1}(k^2+l^2)^{1-m} < \infty$ as soon as $m > 2$. $\diamondsuit$

Remark 4 *More generally, in dimension $N \geq 2$, the same proof shows that: for any $m > 1 + N/2$, $H^m(\Omega) \cap H_0^1(\Omega) \subset \mathcal{L}^1$.*

Let us now give the rate of convergence of the Pure Greedy Algorithm. For the details of the proof, we again refer to [4]. The proof is based on the fundamental lemma:

Lemma 8 ([4, Lemma 3.5]) *Let us assume that $g \in \mathcal{L}^1$. Then, for any $n \geq 0$, $g_n \in \mathcal{L}^1$ and we have:*

$$\|r_{n+1} \otimes s_{n+1}\| = \frac{\langle g_n, r_{n+1} \otimes s_{n+1}\rangle}{\|r_{n+1} \otimes s_{n+1}\|} \geq \frac{\|g_n\|^2}{\|g_n\|_{\mathcal{L}^1}}.$$

The following technical result (easily obtained by induction) is also needed.

Lemma 9 ([4, Lemma 3.4]) *Let $(a_n)_{n\geq 1}$ be a sequence of non-negative real numbers and A a positive real number such that $a_1 \leq A$ and $a_{n+1} \leq a_n \left(1 - \frac{a_n}{A}\right)$. Then, $\forall n \geq 1$,*

$$a_n \leq \frac{A}{n}.$$

Using Lemma 8 and Lemma 9, it is possible to show:

Theorem 3 ([4, Theorem 3.6]) *For $g \in \mathcal{L}^1$, we have*

$$\|g_n\| \leq \|g\|^{2/3} \|g\|_{\mathcal{L}^1}^{1/3} n^{-1/6}. \tag{34}$$

A better rate of convergence can be proven for the Orthogonal Greedy Algorithm. For the Orthogonal Greedy Algorithm, the following Lemma plays the role of Lemma 8.

Lemma 10 *Assume that $g \in \mathcal{L}^1$. Then, for any $n \geq 0$, $g_n^o \in \mathcal{L}^1$ and we have:*

$$\|r_{n+1}^o \otimes s_{n+1}^o\| = \frac{\langle g_n^o, r_{n+1}^o \otimes s_{n+1}^o\rangle}{\|r_{n+1}^o \otimes s_{n+1}^o\|} \geq \frac{\|g_n^o\|^2}{\|g\|_{\mathcal{L}^1}}.$$

Proof. Since $g_n = g - \sum_{k=1}^n \alpha_k r_k \otimes s_k$, it is clear that $g_n \in \mathcal{L}^1$. The equality $\|r_{n+1}^o \otimes s_{n+1}^o\| = \frac{\langle g_n^o, r_{n+1}^o \otimes s_{n+1}^o \rangle}{\|r_{n+1}^o \otimes s_{n+1}^o\|}$ is obtained as a consequence of the Euler-Lagrange equations associated to the optimization problem on (r_{n+1}^o, s_{n+1}^o) (see (19)).

Since $g \in \mathcal{L}^1$, for any $\varepsilon > 0$, we can write $g = \sum_{k\geq 0} c_k u_k \otimes v_k$ with $\|u_k \otimes v_k\| = 1$, and $\sum_{k\geq 0} |c_k| \leq \|g\|_{\mathcal{L}^1} + \varepsilon$. By (9), we have $\langle g - g_n^o, g_n^o\rangle = 0$, and therefore, using Lemma 6:

$$\begin{aligned}
\|g_n^o\|^2 &= \langle g_n^o, g\rangle \\
&= \left\langle g_n^o, \sum_{k\geq 0} c_k u_k \otimes v_k \right\rangle \\
&= \sum_{k\geq 0} c_k \langle g_n^o, u_k \otimes v_k\rangle \\
&\leq \sum_{k\geq 0} |c_k| \frac{\langle g_n^o, r_{n+1}^o \otimes s_{n+1}^o\rangle}{\|r_{n+1}^o \otimes s_{n+1}^o\|} \\
&= (\|g\|_{\mathcal{L}^1} + \varepsilon) \frac{\langle g_n^o, r_{n+1}^o \otimes s_{n+1}^o\rangle}{\|r_{n+1}^o \otimes s_{n+1}^o\|},
\end{aligned}$$

from which we conclude letting ε vanish. $\diamondsuit$

Theorem 4 ([4, Theorem 3.7]) *For $g \in \mathcal{L}^1$, we have*

$$\|g_n^o\| \leq \|g\|_{\mathcal{L}^1}\, n^{-1/2}. \tag{35}$$

Proof. We have, using (16) and Lemma 10:

$$\begin{aligned}
\|g_{n+1}^o\|^2 &= \left\| g - \sum_{k=1}^{n+1} \alpha_k r_k^o \otimes s_k^o \right\|^2 \\
&\leq \|g_n^o - r_{n+1}^o \otimes s_{n+1}^o\|^2 \\
&= \|g_n^o\|^2 - \|r_{n+1}^o \otimes s_{n+1}^o\|^2 \\
&= \|g_n^o\|^2 \left(1 - \frac{\|r_{n+1}^o \otimes s_{n+1}^o\|^2}{\|g_n^o\|^2}\right) \\
&\leq \|g_n^o\|^2 \left(1 - \frac{\|g_n^o\|^2}{\|g\|_{\mathcal{L}^1}^2}\right).
\end{aligned}$$

The conclusion is reached applying Lemma 9 with $a_n = \|g_{n-1}^o\|^2$ and $A = \|g\|_{\mathcal{L}^1}^2$. $\diamondsuit$

Remark 5 *The rate of convergence of the Pure Greedy Algorithm in* (34) *may be improved to $n^{-11/62}$ [6]. For both algorithms, it is known that there exists dictionaries and right-hand sides f (even simple ones, like a sum of only two elements of the dictionary) such that the rate of convergence $n^{-1/2}$ is attained (see [7, 4, 2]). In that sense, the Orthogonal Greedy Algorithm realizes the optimal rate of convergence. Notice that this rate of convergence does not depend on the dimension of the problem. However, the assumption $g \in \mathcal{L}^1$ seems to be more and more demanding, in terms of regularity, as the dimension increases (see Remark 4).*

4 Discussion and open problems

We begin this section by considering the case when the Laplace operator is replaced by the identity operator. We examine on this simplified case the discrepancy between the variational approach consisting in minimizing the energy and the non variational approach solving the Euler-Lagrange equation.

4.1 The Singular Value Decomposition case

The algorithms we have presented above are closely related to the Singular Value Decomposition (SVD, also called *rank one decomposition*). More precisely, omitting the gradient in the optimization problem (7) yields: find $r_n \in L^2(\Omega_x)$ and $s_n \in L^2(\Omega_y)$ such that

$$(r_n, s_n) = \arg \min_{(r,s)\in L^2(\Omega_x)\times L^2(\Omega_y)} \int_\Omega |g_{n-1} - r \otimes s|^2, \tag{36}$$

with the recursion relation

$$g_n = g_{n-1} - r_n \otimes s_n,$$

and $g_0 = g$.

In view of the exact same arguments as in the previous sections, the series $\sum_{n\geq 1} r_n \otimes s_n$ can be shown to converge to g in $L^2(\Omega)$. This problem has a well-known companion discrete problem, namely the SVD decomposition of a matrix (see for example [9]). This corresponds to the case $\Omega_x = \{1, \dots, p\}$, $\Omega_y = \{1, \dots, q\}$, the integral $\int_\Omega$ is replaced by the discrete sum $\sum_{(i,j)\in 1,\dots,p\times 1,\dots,q}$, G is a matrix in $\mathbb{R}^{p\times q}$ and (R_n, S_n) are two (column) vectors in $\mathbb{R}^p \times \mathbb{R}^q$. In this case the tensor product $R_n \otimes S_n$ is simply the matrix $R_n(S_n)^T$. The matrices $G_n \in \mathbb{R}^{p\times q}$ are then defined by recursion: $G_0 = G$ and $G_n = G_{n-1} - R_n(S_n)^T$.

4.1.1 Orthogonality property

An important property of the sequence (r_n, s_n) generated by the algorithm in the SVD case is the orthogonality relation: if $n \neq m$

$$\int_{\Omega_x} r_n r_m = \int_{\Omega_y} s_n s_m = 0. \tag{37}$$

In order to check this, let us first write the Euler-Lagrange equations in the SVD case (compare with (14)): for any functions $(r, s) \in L^2(\Omega_x) \times L^2(\Omega_y)$,

$$\int_\Omega r_n \otimes s_n (r_n \otimes s + r \otimes s_n) = \int_\Omega g_{n-1}(r_n \otimes s + r \otimes s_n). \tag{38}$$

This also reads (compare with (15)):

$$\begin{cases} \left(\int_{\Omega_y} |s_n|^2\right) r_n = \int_{\Omega_y} g_{n-1}\, s_n, \\ \\ \left(\int_{\Omega_x} |r_n|^2\right) s_n = \int_{\Omega_x} g_{n-1}\, r_n. \end{cases} \tag{39}$$

It is immediate to see that (38) for $n = 1$ and $n = 2$ implies,

$$\int_\Omega (r_2 \otimes s_2)(r_2 \otimes s_1) = \int_\Omega (r_2 \otimes s_2)(r_1 \otimes s_2) = 0.$$

Likewise, it can be shown, for any $n \geq 2$ and any $l \in \{2, \dots n\}$

$$\int_\Omega \sum_{k=l}^{n} (r_k \otimes s_k)\,(r_n \otimes s_{l-1}) = \int_\Omega \sum_{k=l}^{n} (r_k \otimes s_k)\,(r_{l-1} \otimes s_n) = 0. \tag{40}$$

The orthogonality property (37) is then easy to check using the Fubini Theorem and arguing by induction.

Remark 6 *A simple consequence of the orthogonality of the functions obtained by the algorithm is that, in the discrete version (SVD of a matrix $G \in \mathbb{R}^{p\times q}$) the algorithm converges in a finite number of iterations (namely* $\max(p,q)$*). As usual in this situation, practice may significantly deviate from the above theory if round-off errors due to floating-point computations are taken into account. This is especially true if the matrix is ill conditioned.*

4.1.2 Consequences of the orthogonality property

The orthogonality property has several consequences: Assume the SVD to be nondegenerate in the sense

$$g = \sum_{n\geq 1} \lambda_n \, u_n \otimes v_n, \tag{41}$$

with

$$\int_{\Omega_x} u_n u_m = \int_{\Omega_y} v_n v_m = \delta_{n,m}, \forall n,m, \text{ and } (\lambda_n)_{n\geq 1} \text{ positive, strictly decreasing,} \tag{42}$$

where $\delta_{n,m}$ is the Kronecker symbol. Then

- (i) The Pure Greedy Algorithm and the Orthogonal Greedy Algorithm are equivalent to one another in the SVD case.
- (ii) The SVD decomposition $g = \sum_{n\geq 1} r_n \otimes s_n$ is unique.
- (iii) At iteration n, $\sum_{k=1}^{n} r_k \otimes s_k$ is the minimizer of $\int_\Omega |g - \sum_{k=1}^{n} \phi_k \otimes \psi_k|^2$ over all possible $(\phi_k, \psi_k)_{1\leq k\leq n} \in \left(L^2(\Omega_x) \times L^2(\Omega_y)\right)^n$.

In addition, simple arguments show that,

- (iv) The only solutions to the Euler Lagrange equations (38) are the null solution $(0,0)$ and the tensor products $\lambda_n u_n \otimes v_n$ (for all $n \geq 1$) in the SVD decomposition of g.
- (v) The solutions to the Euler-Lagrange equations which maximize the L^2-norm $\left(\int_\Omega |r\otimes s|^2\right)^{1/2}$ are exactly the solutions to the variational problem (36).
- (vi) In dimension $N = 2$, the solutions to the Euler-Lagrange equation that satisfy the second order optimality conditions are exactly the solutions of the original variational problem (36).

Notice that there is no loss of generality in assuming $\lambda_n > 0$, and $(\lambda_n)_{n\geq 1}$ decreasing in (41) (up to a change of the (u_n, v_n)). The fundamental assumption in nondegeneracy is thus that $\lambda_n \neq \lambda_m$ if $n \neq m$. When the decomposition has some degeneracy (*i.e.* several n correspond to the same λ_n in (41)) then properties (i)-(iii)-(v)-(vi) still hold true. On the other hand, in (ii) the SVD is only unique up to rotations within eigenspaces and property (iv) must be modified accordingly. In short, the only other solutions beyond those mentioned above consist of tensor products of linear combinations of functions within a given eigenspace. We skip such technicalities. The degenerate case indeed does not differ much from the non degenerate case above in the sense that a complete understanding of the algorithm, both in its variational and in its non variational forms, is at hand.

Let us briefly outline the proofs of assertions (iv)-(v)-(vi).

We first prove assertion (iv). It is sufficient to consider the first iteration of the algorithm. Using the SVD decomposition of g, the Euler-Lagrange equations write: for any functions $(r,s) \in L^2(\Omega_x) \times L^2(\Omega_y)$,

$$\int_\Omega r_1 \otimes s_1 (r_1 \otimes s + r \otimes s_1) = \sum_{n \geq 1} \lambda_n \int_\Omega u_n \otimes v_n (r \otimes s_1 + r_1 \otimes s).$$

Using the orthogonality property, and successively $(r,s) = (0, v_n)$ and $(r,s) = (u_n, 0)$ as test functions, we get

$$\begin{cases} \displaystyle\int_{\Omega_x} |r_1|^2 \int_{\Omega_y} s_1 v_n = \lambda_n \int_{\Omega_x} r_1 u_n, \\ \displaystyle\int_{\Omega_y} |s_1|^2 \int_{\Omega_x} r_1 u_n = \lambda_n \int_{\Omega_y} s_1 v_n, \end{cases}$$

which yields: $\forall n \geq 1$

$$\int_{\Omega_y} s_1 v_n \int_{\Omega_x} r_1 u_n \left(\int_{\Omega_x} |r_1|^2 \int_{\Omega_y} |s_1|^2 - (\lambda_n)^2 \right) = 0.$$

Since for $n \neq m$, $\lambda_n \neq \lambda_m$, this shows that either $r_1 \otimes s_1 = 0$, or there exists a unique n_0 such that $\lambda_{n_0} = \sqrt{\int_\Omega |r_1 \otimes s_1|^2}$ and $\forall n \neq n_0$, $\int_{\Omega_y} s_1 v_n = \int_{\Omega_x} r_1 u_n = 0$ (because the product $\int_{\Omega_y} s_1 v_n \int_{\Omega_x} r_1 u_n$ cancels and thus each of the term cancels because of the Euler Lagrange equations). Since by the Euler-Lagrange equations, r_1 (resp. s_1) can be decomposed on the set of orthogonal functions $(u_n,\, n \geq 1)$ (resp. $(v_n,\, n \geq 1)$), we get $r_1 \otimes s_1 = \lambda_{n_0} u_{n_0} \otimes v_{n_0}$, which concludes the proof of assertion (iv). Assertion (v) is readily obtained using (iv) and the orthogonality property. Notice that assertion (ii) is a consequence of assertions (iv)-(v). To prove assertion (vi), we recall that the second order optimality condition writes (see Lemma 5, adapted to the SVD case): $\forall (r,s) \in L^2(\Omega_x) \times L^2(\Omega_y)$,

$$\left(\int_\Omega (r_n \otimes s_n - g_n) r \otimes s \right)^2 \leq \int_\Omega |r \otimes s_n|^2 \int_\Omega |r_n \otimes s|^2. \tag{43}$$

It is again enough to consider the case $n = 1$. Let us consider a solution of the Euler-Lagrange equation: $r_1 \otimes s_1 = \lambda_{n_0} u_{n_0} \otimes v_{n_0}$, and let us take as test functions in (43) $(r,s) = (u_n, v_n)$, for all $n \geq 1$. We obtain that for all $n \geq 1$, $(\lambda_n)^2 \leq (\lambda_{n_0})^2$ which concludes the proof of assertion (vi). Notice that in dimension $N \geq 3$, assertion (vi) seemingly does not hold: the solutions to the Euler-Lagrange equation that satisfy the second order optimality conditions may not necessarily be global minimizers.

4.1.3 Link between the Euler-Lagrange equations and the variational problem

Properties (iv)-(v)-(vi) above tend to indicate that, at least in the SVD case, the consideration of the solutions to the Euler-Lagrange equations is somehow close to the consideration of the minimization problems.

Indeed, if we assume that at each iteration, non zero solutions of the Euler-Lagrange equations are obtained (of course under the assumption $g_{n-1} \neq 0$ in (38)), then the non variational form of the algorithm, if it converges, will eventually provide the correct decomposition. We however would like to mention two practical difficulties.

First, it is not clear in practice how to compute the norm $\|g_n\|$ to check the convergence, since this is in general a high dimensional integral. A more realistic convergence criterion

would read: $\|r_n \otimes s_n\|$ *is small compared to* $\left\|\sum_{k=1}^{n-1} r_k \otimes s_k\right\|$. However, using this criterion, it is possible to erroneously conclude that the algorithm has converged, while a term with an arbitrarily large contribution has been missed. Indeed, consider again, to convey the idea, the case (41)-(42). Assume that the tensor product $\lambda_2 u_2 \otimes v_2$ is picked at first iteration (*instead of* the tensor product $\lambda_1 u_1 \otimes v_1$ which would be selected by the *variational* version of the algorithm). Assume similarly that $\lambda_3 u_3 \otimes v_3$ is picked at second iteration, and so on and so forth. In such a situation, one would then decide the series $\sum_{n \geq 2} \lambda_n u_n \otimes v_n$ solves the problem, while obsviously it does not. We will show below (see Section 4.1.4) that in the simple fixed-point procedure we have described above to solve the nonlinear Euler-Lagrange equations, the fact that $\lambda_1 u_1 \otimes v_1$ is missed, and never obtained as a solution, may indeed happen as soon as the initial condition of the iterative procedure has a zero component on the eigenspace associated to λ_1.

Second, without an additional assumption reminiscent of the minimizing character of the solution, iteratively solving the Euler-Lagrange equations may result in picking the tensor products $\lambda_n u_n \otimes v_n$ in an order not appropriate for computational efficiency. Such an assumption is present in assertions (v) and (vi). For the illustration, let us indeed consider a SVD decomposition

$$g = \sum_{n \geq 1} \lambda_n u_n \otimes v_n$$

for some functions u_n and v_n that become highly oscillatory when n grows. It is clear that we may obtain an error in H^1 norm that is arbitrarily large at each iteration of the algorithm. In particular, it may happen (in particular if smooth functions are chosen as initial guesses for the nonlinear iteration loop solving the Euler-Lagrange equation) that the highly oscillatory products are only selected in the latest iterations, although they contribute to the error in a major way. A poor efficiency of the algorithm follows. Inevitably, reaching computational efficiency therefore requires to account for some additional assumptions to select the appropriate solutions among the many solutions of the Euler-Lagrange equations.

In the spirit of the above discussion, one can notice that

- (vii) The null solution $(0, 0)$ to the Euler-Lagrange equation (38) is generically not isolated within the set of all solutions.

Indeed, consider a SVD $g = \sum_{n \geq 1} \lambda_n \, u_n \otimes v_n$, such that u_n and v_n are non-zero functions for all $n \geq 1$ (and $\lambda_n > 0$). Then, any $(\lambda_n u_n, v_n)$ is a solution of the Euler Lagrange equation at the first iteration, and the norm of the $(\lambda_n u_n, v_n)$ which is selected may be arbitrarily small since the series $\sum_{n \geq 1} \lambda_n \, u_n \otimes v_n$ converges, and therefore $\|\lambda_n u_n \otimes v_n\|$ goes to zero. A similar argument applies to all iterations of the algorithm. Therefore, a criterion of convergence of the type $\|r_n \otimes s_n\|$ *is small compared to* $\left\|\sum_{k=1}^{n-1} r_k \otimes s_k\right\|$ may again yield an erroneous conclusion and lead to a prematurate termination of the iterations.

Remark 7 *Note of course that the relaxation step performed in the orthogonal version of the algorithm does not solve any of the above difficulties.*

4.1.4 Resolution of the Euler-Lagrange equations

A last comment we would like to make on the SVD case again concerns the practical implementation of the solution procedure for the Euler-Lagrange equations. Consider the

discrete case for clarity. The fixed point procedure then simply writes (for a fixed n): at iteration $k \geq 0$, compute two vectors $(R_n^k, S_n^k) \in \mathbb{R}^p \times \mathbb{R}^q$ such that:

$$\begin{cases} (S_n^k)^T S_n^k R_n^{k+1} = G_{n-1} S_n^k, \\ (R_n^{k+1})^T R_n^{k+1} S_n^{k+1} = (G_{n-1})^T R_n^{k+1}. \end{cases} \tag{44}$$

One can check that this procedure is similar to the power method to compute the largest eigenvalues (and associated eigenvectors) of the matrix $(G_{n-1})^T G_{n-1}$. Let us explain this. The recursion writes:

$$S^{k+1} = (G^T G) S^k \frac{\|S^k\|^2}{\|GS^k\|^2},$$

where $\|\cdot\|$ here denotes the Euclidean norm and where we have omitted the subscripts n and $n-1$ for clarity. To study the convergence of this algorithm one can assume that G is actually a diagonal matrix up to a change of coordinate. Indeed, let us introduce the SVD decomposition of G: $G = U\Sigma V^T$ where U and V are two orthogonal matrices, and Σ is a diagonal matrix with non-negative coefficients. Without loss of generality, we may assume that $q \leq p$, $U \in \mathbb{R}^{p\times q}$, $\Sigma \in \mathbb{R}^{q\times q}$, $V \in \mathbb{R}^{q\times q}$ and $\Sigma_{1,1} \geq \Sigma_{2,2} \geq \ldots \geq \Sigma_{q,q}$. For simplicity, assume that $\Sigma_{1,1} > \Sigma_{2,2} > 0$. Then, setting $\tilde{S}^k = V^T S^k$, the recursion reads $\tilde{S}^{k+1} = (\Sigma^T\Sigma)\tilde{S}^k \frac{\|\tilde{S}^k\|^2}{\|\Sigma\tilde{S}^k\|^2}$ and the convergence is easy to study. One can check that if the initial condition S^0 has a non-zero component along the vector associated to the largest value $\Sigma_{1,1}$, then S^k converges to this vector. The convergence is geometric, with a rate related to $\frac{\Sigma_{2,2}}{\Sigma_{1,1}}$ (at least if the initial condition S^0 has a non-zero component along the vector associated to $\Sigma_{2,2}$, otherwise $\Sigma_{2,2}$ should be replaced by the appropriate largest $\Sigma_{k,k}$, with $k > 1$). Of course, if the initial condition is not well chosen (namely, if S^0 has a zero component along the vector associated to $\Sigma_{1,1}$), then this algorithm cannot converge to the solution of the variational version of the algorithm.

We would like to mention that this method to compute the SVD of a matrix is actually known to poorly perform in practice. More precisely, the approach is very sensitive to numerical perturbations, see [9, Lecture 31]) since the condition number of $(G_{n-1})^T G_{n-1}$ is typically large. Alternative methods exist that compute the SVD decomposition, and it would be interesting to use these techniques as guidelines to build more efficient procedures to solve the nonlinear Euler-Lagrange equations (15).

4.2 Euler-Lagrange approach for the Poisson problem

We now return to the solution of the Poisson problem. Our purpose is to see which of the above mentioned difficulties survive in this case. We shall also see new difficulties appear.

We first observe, on a general note, that a property similar to (40) holds in the Poisson case, namely:

$$\int_\Omega \nabla\left(\sum_{k=l}^n r_k \otimes s_k\right) \cdot \nabla(r_n \otimes s_{l-1}) = \int_\Omega \nabla\left(\sum_{k=l}^n r_k \otimes s_k\right) \cdot \nabla(r_{l-1} \otimes s_n) = 0. \tag{45}$$

This, however, does not seem to imply any simple orthogonality property as (37). In particular, in the Poisson case, it is generally wrong that, for $n \neq m$, $\int_{\Omega_x} \nabla(r_n \otimes s_n) \cdot \nabla(r_m \otimes s_m) = 0$.

Next, we remark that none of the properties (i)-(ii)-(iii) holds in the Poisson case. Likewise, we are not able to characterize the list of solutions to the Euler-Lagrange equations as we did in (iv)-(v)-(vi).

This is for the generic situation, but in order to better demonstrate the connections between the SVD case above and the Poisson case, let us show that, in fact, the Poisson case necessarily embeds all the difficulties of the SVD case. For this purpose, we consider the original algorithm (for the Poisson problem) performed for a particular right-hand-side $f = -\Delta g$, namely

$$\begin{cases} g = \sum_{k=1}^{N} \alpha_k \phi_k \otimes \psi_k \text{ where } \alpha_k \in \mathbb{R}, \\ \phi_k \text{ (resp. } \psi_k\text{) are eigenfunctions of} \\ \text{the homogeneous Dirichlet operator } -\partial_{xx} \text{ (resp. } -\partial_{yy}\text{)} \\ \text{and satisfy } \forall k,l,\ \int \phi_k \phi_l = \int \psi_k \psi_l = \delta_{k,l}, \end{cases} \tag{46}$$

where $\delta_{k,l}$ is again the Kronecker symbol. Then, it can be shown that, as in the SVD case, $r_k \otimes s_k = \alpha_k \phi_k \otimes \psi_k$ are indeed solution to the Euler-Lagrange equations (14). This suffices to show the non uniqueness of the solution. Furthermore, and in sharp contrast to (iv), there even exist solutions to the Euler Lagrange equations that are not of the above form.

Here is an example of the latter claim. Consider the case $\phi_1 = \psi_1$, associated with an eigenvalue λ_1 and $\phi_2 = \psi_2$, associated with an eigenvalue $\lambda_2 \neq \lambda_1$. We suppose $\alpha_k = 0$ for $k \geq 2$. We are looking for r and s solution to the Euler-Lagrange equations

$$\begin{cases} -\displaystyle\int |s|^2 r'' + \int |s'|^2 r = \int fs, \\ \\ -\displaystyle\int |r|^2 s'' + \int |r'|^2 s = \int fr. \end{cases}$$

Then, it can be checked that $r = r_1\phi_1 + r_2\phi_2$ and $s = s_1\psi_1 + s_2\psi_2$ are solution to the Euler-Lagrange equations, with the following set of parameters: $r_1 = 1$, $r_2 = 1/2$, $s_1 = 2$, $s_2 = 1$, $\alpha_1 = \frac{9\lambda_1+\lambda_2}{4\lambda_1}$ and $\alpha_2 = \frac{2\lambda_1+3\lambda_2}{2\lambda_2}$. Likewise, it is immediate to see that (vii) still holds. In view of the above remarks, it seems difficult to devise (and, even more difficult, to prove the convergence of) efficient iterative procedures to correctly solve the Euler-Lagrange equation.

4.3 Some numerical experiments and the non self-adjoint case

We now show some numerical tests. Even though the algorithms presented above have been designed for solving problems in high dimension, we restrict ourselves to the two-dimensional case. For numerical results in higher dimension, we refer to [1]. Moreover, we consider the discrete case mentioned in Section 4.1, which writes (compare with (3)): for a given symmetric positive definite matrix $D \in \mathbb{R}^{d\times d}$ (which plays the role of the one-dimensional operator $-\partial_{xx}$), and a given matrix $F \in \mathbb{R}^{d\times d}$ (which plays the role of the right-hand side f):

$$\text{Find } G \in \mathbb{R}^{d\times d} \text{ such that } DG + GD = F. \tag{47}$$

Here, the dimension d typically corresponds to the number of points used to discretize the one-dimensional functions r_n or s_n. To this problem is associated the variational problem (compare with (4))

$$\text{Find } G \in \mathbb{R}^{d\times d} \text{ such that } G = \arg\min_{U\in\mathbb{R}^{d\times d}} \left(\frac{DU+UD}{2} - F\right) : U, \tag{48}$$

where, for two matrices $A, B \in \mathbb{R}^{d\times d}$, $A : B = \sum_{1\leq i,j\leq d} A_{i,j}B_{i,j}$. The matrix G is built as a sum of rank one matrices $R_k S_k^T$ with $(R_k, S_k) \in (\mathbb{R}^d)^2$, using the following Pure Greedy Algorithm (compare with the algorithm presented in Section 2.1): `Set` $F_0 = F$ `and at iteration` $n \geq 1$,

1. Find R_n and S_n two vectors in $\mathbb{R}^d$ such that:

$$(R_n, S_n) = \arg \min_{(R,S)\in(\mathbb{R}^d)^2} \left(\frac{D(RS^T) + (RS^T)D}{2} - F_{n-1} \right) : (RS^T). \qquad (49)$$

2. Set[1] $F_n = F_{n-1} - (DR_nS_n^T + R_nS_n^TD)$.

3. If $\|F_n\| > \varepsilon$, proceed to iteration $n+1$. Otherwise stop.

As explained in Section 2.3, Step 1 of the above algorithm is replaced in practice by the resolution of the associated Euler-Lagrange equations. This consists in finding two vectors R_n and S_n in $\mathbb{R}^d$ solution to the nonlinear equations:

$$\left\{ \begin{array}{l} \|S_n\|^2 \, DR_n + \|S_n\|_D^2 \, R_n = F_{n-1}S_n, \\ \|R_n\|^2 \, DS_n + \|R_n\|_D^2 \, S_n = F_{n-1}^T R_n, \end{array} \right. \qquad (50)$$

where, for any vectors $R \in \mathbb{R}^d$, we set $\|R\|_D^2 = R^TDR$. This nonlinear problem is solved by a simple fixed point procedure (as (24)). We have observed in practice that choosing a random vector as an initial condition for the fixed point procedure is more efficient than taking a given deterministic vector (like $(1,\ldots,1)^T$). This is of course related to the convergence properties of the fixed point procedure we discussed in Section 4.1.4.

4.3.1 Convergence of the method

In this section, we take D diagonal, with $(1,2,\ldots,d)$ on the diagonal, and a random matrix F. The parameter ε is 10^{-6}. We observe that the algorithm always converges. This means that, in practice, the solutions of the Euler-Lagrange equations (50) selected by the fixed point procedure are appropriate.

On Figure 1, we plot the energy $\left(\frac{DU_n+U_nD}{2} - F\right) : U_n$, where $U_n = \sum_{k=1}^n R_kS_k^T$. We observe that the energy rapidly decreases and next reaches a plateau. This is a general feature that we observe on all the tests we perfomed.

In Table 1, we give the number of iterations necessary for convergence, as a function of d. We observe a linear dependency, which unfortunately we are unable to explain theoretically.

d	10	20	30
Number of iterations	22-23	45-46	69-70

Table 1: Number of iterations typically needed for convergence as a function of d, for various random matrices F ($D = diag([linspace(1,2,d)])$, $\varepsilon = 10^{-6}$).

4.3.2 The non self-adjoint case

In [1], it is actually proposed to use the Orthogonal Greedy Algorithm for non self adjoint operators.

Consider, for the prototypical case of an advection diffusion equation:

$$\text{Find } g \in H_0^1(\Omega) \text{ such that } \left\{ \begin{array}{rl} a\cdot\nabla g - \Delta g = f & \text{in } \Omega, \\ g = 0 & \text{on } \partial\Omega, \end{array} \right. \qquad (51)$$

[1] In practice, to avoid numerical cancellation, we actually set $F_n = F - (DU_n + U_nD)$ where $U_n = \sum_{k=1}^n R_kS_k^T$.

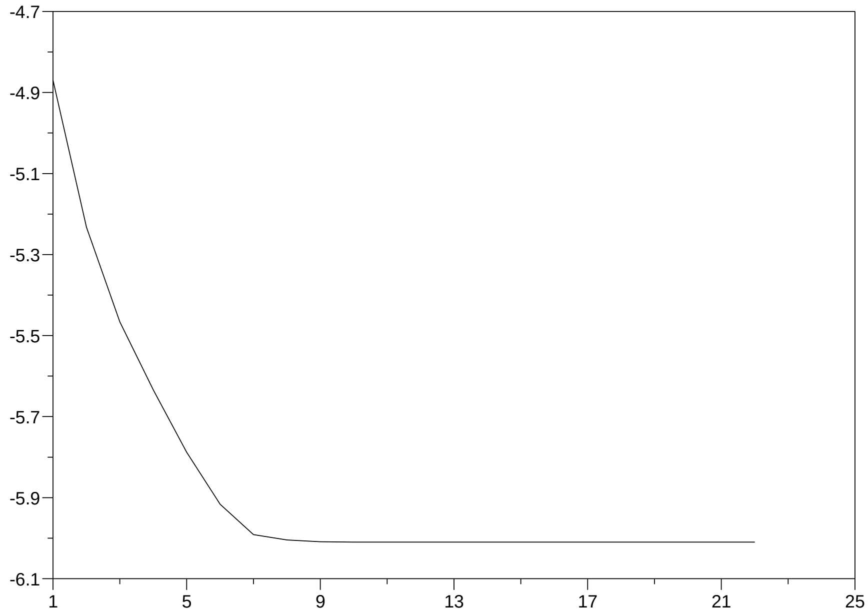

Figure 1: Evolution of the energy as a function of iterations ($d = 10$, $D = diag([linspace(1,2,d)])$, $\varepsilon = 10^{-6}$).

where $a : \Omega \to \mathbb{R}^2$ is a given smooth velocity field. When $a = \nabla V$ for some real-valued function V, problem (51) is equivalent to minimizing the energy

$$\frac{1}{2}\int_\Omega |\nabla u|^2 \exp(-V) - \int fu \exp(-V).$$

When this is not the case, it is not in general possible to recast (51) in terms of a minimization problem. However, a variational formulation can be written as: Find $g \in H_0^1(\Omega)$ such that, for all $v \in H_0^1(\Omega)$,

$$\int_\Omega (a \cdot \nabla g)v + \nabla g \cdot \nabla v = \int_\Omega fv.$$

It is proposed in [1] to use this variational formulation in step 1 and 2 of the Orthogonal Greedy Algorithm. The iterations then write: `set` $f_0 = f$, `and at iteration` $n \geq 1$,

1. `Find` $r_n \in H_0^1(\Omega_x)$ `and` $s_n \in H_0^1(\Omega_y)$ `such that, for all functions` $(r,s) \in H_0^1(\Omega_x) \times H_0^1(\Omega_y)$,

$$\int_\Omega (a\cdot\nabla(r_n\otimes s_n))(r_n\otimes s+r\otimes s_n)+\nabla(r_n\otimes s_n)\cdot\nabla(r_n\otimes s+r\otimes s_n) = \int_\Omega f_{n-1}(r_n\otimes s+r\otimes s_n). \tag{52}$$

2. `Find` $u_n \in \mathrm{Vect}(r_1 \otimes s_1, \ldots, r_n \otimes s_n)$ `such that for all` $v \in \mathrm{Vect}(r_1 \otimes s_1, \ldots, r_n \otimes s_n)$

$$\int_\Omega (a \cdot \nabla u_n)v + \nabla u_n \cdot \nabla v = \int_\Omega fv. \tag{53}$$

3. `Set` $f_n = f_{n-1} - (a \cdot \nabla u_n - \Delta u_n)$.

4. `If` $\|f_n\|_{H^{-1}(\Omega)} \geq \varepsilon$, `proceed to iteration` $n+1$. `Otherwise, stop.`

The corresponding discrete formulation reads:

$$\text{Find } G \in \mathbb{R}^{d \times d} \text{ such that } BG + GB^T = F, \tag{54}$$

where B is not supposed to be symmetric here (compare to (47)). The numerical method reads: `Set` $F_0 = F$ `and at iteration` $n \geq 1$,

1. `Find` R_n `and` S_n `two vectors in` $\mathbb{R}^d$ `such that:`

$$\begin{cases} \|S_n\|^2 BR_n + \|S_n\|_B^2 R_n = F_{n-1}S_n, \\ \|R_n\|^2 BS_n + \|R_n\|_B^2 S_n = F_{n-1}^T R_n. \end{cases} \tag{55}$$

2. `Set` $F_n = F_{n-1} - (BR_nS_n^T + R_nS_n^TB^T)$.

3. `If` $\|F_n\| > \varepsilon$, `proceed to iteration` $n+1$. `Otherwise stop.`

We consider the case when $B = D + A$ with D symmetric positive definite, and A antisymmetric, so that we know there exists a unique solution to (54). On the numerical tests we have performed, the algorithm seems to converge. In the absence of any energy minimization principle, it is however unclear to us how to prove convergence of this algorithm.

References

[1] A. Ammar, B. Mokdad, F. Chinesta, and R. Keunings. A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modeling of complex fluids. *J. Non-Newtonian Fluid Mech.*, 139:153–176, 2006.

[2] A.R. Barron, A. Cohen, W. Dahmen, and R.A. DeVore. Approximation and learning by greedy algorithms. *Annals of Statistics*, 36(1):64–94, 2008.

[3] G. Davis, S. Mallat, and M. Avellaneda. Adaptive greedy approximations. *Constr. Approx.*, 13(1):57–98, 1997.

[4] R.A. DeVore and V.N. Temlyakov. Some remarks on greedy algorithms. *Adv. Comput. Math.*, 5:173–187, 1996.

[5] L.K. Jones. On a conjecture of Huber concerning the convergence of projection pursuit regression. *Ann. Stat.*, 15(2):880–882, 1987.

[6] S.V. Konyagin and V.N. Temlyakov. Rate of convergence of pure greedy algorithm. *East J. Approx.*, 5:493–499, 1999.

[7] E.D. Livshitz and V.N. Temlyakov. Two lower estimates in greedy approximation. *Constr. Approx.*, 19:509–524, 2003.

[8] V.N. Temlyakov. Greedy approximation. *Acta Numerica*, 17:235–409, 2008.

[9] L.N. Trefethen and D. Bau. *Numerical linear algebra*. SIAM, 1997.